

Figure 1

Schematic layout of the arrangement of the genetic locus encoding the signal peptide precursor, the histidine kinase and the response regulator. Note that this arrangement is different from other loci in related streptococci for the following reasons: a) The comC gene is transcribed from its own promoter alone, unlike the genes thus far described in other streptococci that are arranged in an operon-like cluster with the comC/DE genes being transcribed from a single promoter.

b) The *comC* gene is separated from the *comD* gene by 148 nucleotides.

Streptococcus mutans *ComCDE Operon*



Figure 2

Sequences of the open reading frames encoding the signal peptide precursor (ComC), the histidine kinase (ComD), and the response regulator (ComE).

➤ *S. mutans* comC gene

Encodes a precursor to a signal peptide

[ATGAAAAAAACACTATCATTAAGGAAATTAAAGACTGATGAATTAG
AGATTATCATTGGCGGA (AGCGGAAGCCTATCAACATTTCGGCTGTTAACAGAAG
TTTACACAAGCTTGGGAAAA)] TAA

➤ *S. mutans* CSP encoding sequence

Competence Signal Peptide

AGCGGAAGCCTATCAACATTTCGGCTGTTAACAGAAGTTTACACAAGCTTGGG
AAAA [SEQ ID NO:1]

➤ *S. mutans* comD gene

Encodes a protein that functions as a histidine kinase receptor

[ATGAATGAAGCCTTAATGATACTTCAAATGGTTATTAACTTATCTAACCGTTCTAT
TTCTCTTGTCTATTCTAAGGTAAGTAATGTCACTTTATCGAAAAAGGAATTAACT
CTTTTTCGATAAGCAATTCTGATAATGATTGCTGTTACGATGGTAACGTAAACCT
GTTTATCCTGCAGAGCCTTTATTAGCTTATCAATTATCTTAATAGACAGA
ATAGTCTTCTCTAAATATAATTATGGTCTGCTGCCAGTCTGACTTGTGTT
AGGCGGGCAATCATATTCTTATCTGGATGGAACCTCAAGGAATTGTAATGGCAGTAG
CATTATAACCACCTATATGATCGAGTTGCAGGAATAGCGCTAACCTCTTCTCA
GTGTGTTCAATGTTGATATTGGTCGACTAAAGATAGTTGACCAAGATGAAGGTCAA
AAACGCTTGATTCCAATGAATATTACTATGCTTCTATACTACCTTTAATACAGGTATT
GTATGTTATAGAGAGTTATAATGTGATACCGACTTAAAATTCTGIAATTGTCGTTA
TTGTCTATCTTATTGATTCTGATCTCATTAAAGCCAATATACCAAACAA
AAGGTTCAAAATGAGATAATTGGCACAAAAGGAAGCTCAGATTGAAATATCACCCAGTA
TAGTCAGCAAATAGAAATCTTTACAAGGATATTGAAAGTTCCGCCATGATTATCTGA
ATATTAACTAGCCTCAGATTAGGCATTGAAAATAAGATTAGCTAGTATTGAAAAG
ATTTCACCATCAAATCTTAGAAAAACAGGCACATTGCAGGATAACCGTTATAATAT
CGGCCATCTAGCTAATATTCAAACGATGCTGCAAGGGTATCTTGTGAGCAAATCT
TAGAAGCTCAGAATAAAAGATTGCTGCAATGAGTCAAGTAAAATACAACCTG
CCTGAGATGGAGTTGCTGATTCTGATTCATTACCATACTTTCTATCTTGTGATAATGCCAT
TGAGGCTGCTTCAATCATAAATCCTGAAATTCACTGAGTTAGCCTTTTAAGAAAAATG
GCAGTATAGTCTTATCATTAGAAATTCCACCAAAGAAAAACAAATAGATGTGAGTAA
ATTTTAAAGAAAATTCACAAAGGCTCCAATCGCGGTATTGGTTAGCAAAGGT
GAATCATATTCTGAAACATTATCCAAAACCAGTTACAAACAAGCAATCATCATCATT
TATTCAAGCAACTCTAATAATAAAA] TAG

➤ *S. mutans* comE gene

Encodes a response regulator that activates transcription of a number of genes

[ATGATTCTATTTGTATTGGAAGATGATTACAAACAAGGACGTCTGAAACCA
CCATTGCAGCTATCATGAAAGAAAAATTGGTCTTATAAAGAATTGACTATTTTGGAA
AAACCACAAACAACCTATTGACGCTATCCCTGAAAAGGGCAATCACCAGATTCTTTT
GGATATTGAAATCAAAAAGAGGAAAAGAAAGGACTGGAAGTAGCCAATCAGATTAGAC
AGCATAATCCTAGTCAGTTATTGCTTTGTACGACACATTCTGAGTTATGCCCTC
ACTTTTCAGTATCAGGTATCGCTTGATTTATTGATAAATCTTGAATCCTGAGGA
GTTCTCCCACCGCATTGAATCAGCGCTGTATTATGCTATGGAAAACAGCCAGAAGAATG
GTCAATCAGAGGAACCTTTATTTCATTCACTGAAACTCAGTTCAGGTCCCTTT
GCTGAGATTCTGTATTTGAAACATCTCAACAGCCCATAAGCTCTGCCTTATACTTA
TGATGAACGGATTGAATTCTACGGCAGTATGACTGACATTGTTAAAATGGATAAGAGAC
TTTTCAGTGCCATCGCTTTATTGTCATCCTGCCAATTACCGTATTGATCGG
AAAAAACGCTTGGCCTATTCGAAATAATAAGCTTGCTTATTCACGAACTAAGTT
AACAAAACGTGAGAGCTGTGATTGCTGATCAAAGGAGAGCAAA] TGA

09623047 04-1000

Figure 3

The amino acid sequences of the signal peptide precursor (ComC), the histidine kinase (ComD), and the response regulator (ComE).

> S. mutans ComC protein (CSP Precursor)

MKKTLSLKNDFKEIKTDELEIIIIGGSGSLSTFFRLFNRSFTQALGK

> S. mutans ComD protein (Histidine Kinase)

MNEALMILSNGLTYLTFLFLLFSKVSVNLTKELTLFSISNFLIMIAVTMVNVNL
FYPAEPLYFIALSIYLNQRQNSLSLNIFYGLLPVASSDLFRRAIIFFILDGTQGIVMGSS
IITTYMIEFAGIALSYLFLSVFNVDIGRLKDSLTKMKVKRLLIPMNITMLLYLLIQVL
YVIESYNVIPTLKERKFVVIVYLFLILISFLSQYTKQKVQNEIMAQKEAQIRNITQY
SQQIESLYKDIRSFRHDYLNILTSRLGIENKDASIEKIYHQILEKTGHQLQDTRYNI
GHLANIQMDAVKGILSAKILEAQNKKIAVNVEVSSKIQLPEMELLDFITILSILCDNAI
EAAFESLNPEIQLAFFKKNGSIVFIIONSTKEQIDVSKIFKENYSTKGSNRGIGLAKV
NHILEHYPKTSLOTSNHHHLFKQLLIIK

> S. mutans ComE protein (Response Regulator)

MISIFVLEDDFLQQGRLETTIAAIMKEKNWSYKELTIFGKPQQLIDAYPEKGNHQIFFL
DIEIKKEEKKGLEVANQIRQHNPSAVIVFVTTHSEFMPLTFQYQVSALDFIDKSLNPEE
FSHRIESALYYAMENSQKNGQSEELFIFHSSETQFQVPPAEILYFETSSTAHKLCLYT
DERIEFYGSMTDIVKMDKRLFQCHRSFIVNPANITRIDRKRLAYFRNNKSCLISRTKL
TKLRAVIADQRRAK

DO NOT DESTROY - CONFIDENTIAL

Express Mail No. EL818623952US
Inventor(s): Cvitkovitch et al
Atty Dkt No.: 1889-00401
Contact: C.M. Mintz (713)238-8000
Sheet 5 of 19

Figure 4

The deduced amino acid sequence of the signal peptide precursor in various strains and its predicted cleavage site. The original peptide is expressed as a 46-amino acid peptide that is cleaved after the glycine-glycine residues to generate an active signal peptide.

BM71	CSP	1	MKKTPSILKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK	46
GB14	CSP	1	MKKTLSILKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK	46
H7	CSP	1	MKKTLSILKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK	46
JH1005	CSP	1	MKKTLSILKNDFKEIKTDELEIIIGGSGTLSTFFRLFNRSFTQAA	43
LT11	CSP	1	MKKTLSILKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK	46
NG8	CSP	1	MKKTLSILKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK	46
UAB159	CSP	1	MKKTLSILKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK	46

consensus: 1 MKKTLSQLNDFKEIKTDELEIIIGG SGSLSTFFRLFNRSFTQALGK 46
predicted cleavage site: ^

Figure 5

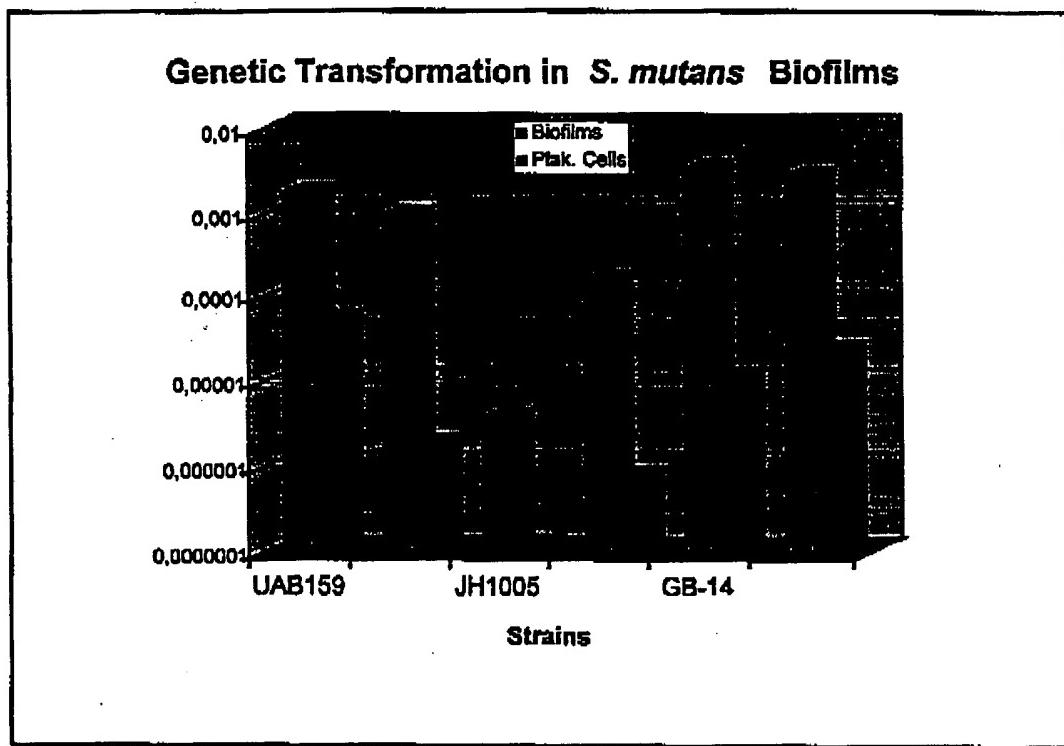
The synthetic signal peptide that is effective at inducing competence, biofilm formation and acid tolerance in *Streptococcus mutans*.

SGSLSTFFRLFNRSFTQALGK [SEQ ID NO:2]

Express Mail No. EL818623952US
Inventor(s): Cvitkovitch et al
Atty Dkt No.: 1889-00401
Contact: C.M. Mintz (713)238-8000
Sheet 6 of 19

Figure 6

The natural activity of the signal/receptor system functioning *in vitro* in model biofilms as determined by the ability of various strains of *S. mutans* to accept donor plasmid DNA conferring erythromycin resistance.



Express Mail No. EL818623952US
Inventor(s): Cvitkovitch et al
Atty Dkt No.: 1889-00401
Contact: C.M. Mintz (713)238-8000
Sheet 7 of 19

Figure 7
Induction of genetic transformation in *Streptococcus mutans*
by synthetic competence stimulating peptide (SCSP)¹

Strain	Peptide added Number of Transformants/Recipients	No peptide Number of Transformants/Recipients
UAB15	4.65×10^{-1}	1.78×10^{-6}
JH1005 ²	6.98×10^{-2}	0

¹The final concentration of SCSP used was 500 ng/ml.

²The strain contains a nonsense mutation in the *comC* gene encoding the CSP.

Figure 8

List of the primers used to amplify the genes or internal regions of the target genes by polymerase chain reaction (PCR) for subsequent sequencing or inactivation.

ComC region

ComC Primer Pair: F5-B5

[F5] 23406-23424 5'- AGTTTTTGTCTGGCTGCG -3'

19 nt forward primer

pct G+C: 47.4 Tm: 50.5

[B5] 24056-24037 5'- TCCACTAAAGGCTCCAATCG -3'

20 nt backward primer

pct G+C: 50.0 Tm: 51.9

651 nt product for F5-B5 pair (23406-24056)

Optimal annealing temp: 50.3

pct G+C: 30.9 Tm: 71.5

ComD region

ComD Primer Pair: F1-B1

[F1] 392-415 5'- CGCTAAGTTACCTCTTCAGTG -3'

24 nt forward primer

pct G+C: 45.8 Tm: 51.6

[B1] 683-663 5'- GCTTCCTTTGTGCCATTATC -3'

21 nt backward primer

pct G+C: 42.9 Tm: 50.8

292 nt product for F1-B1 pair (392-683)

Optimal annealing temp: 49.5

pct G+C: 30.8 Tm: 70.2

ComE region

ComE Primer Pair: F1-B1

[F1] 145-165 5'- CCTGAAAAGGGCAATCACCAAG -3'

21 nt forward primer

pct G+C: 52.4 Tm: 55.9

[B1] 606-585 5'- GCGATGGCACTGAAAAAGTCTC -3'

22 nt backward primer

pct G+C: 50.0 Tm: 55.4

462 nt product for F1-B1 pair (145-606)

Optimal annealing temp: 53.6

pct G+C: 38.3 Tm: 74.1

Figure 9

ComCDE local region. The ComC (first highlighted region; nucleotides 101 to 241), ComD (second highlighted region; nucleotides 383 to 1708) and ComE (third highlighted region; nucleotides 1705 to 2457) proteins are highlighted.

Sequence Range: 1 to 2557

10 20 30 40 50
ACATTATGTGCTCTAAGGAAAATATTACTTTCAAGAAAATCCATGATT
TGTAATAACACAGGATTCTTTATAATGAAAAAGTTCTTTAGGTACTAA
<K K L F I W S K
<

60 70 80 90 100
TTTCATAAAAAAATAGTATACTAATTATAATCAAAAAAAGGAGATATAAA
AAAAGTATTTTATCATATGATTAATATTAGTTTTCTCTATATT
<K M F F L I S I I I L F L L Y L
<

110 120 130 140 150
ATGAAAAAAACACTATCATAAAAAATGACTTTAAAGAAAATTAAAGACTGA
TACTTTTTGTGATAGTAATTTTTACTGAAATTCTTTAATTCTGACT
M E K K T I S V K N D A E K R A S K V A T E D
ORF RF [2]
<I F F V S D N F F S K L S I L V S
<

160 170 180 190 200
TGAATTAGAGATTATCATGGCGGAAGCGGAAGCCTATCAACATTTTC
ACTTAATCTCTAATAGTAACCGCCTTCGCCCTCGGATAGTTGTAAAAAGG
E V E E L I Y I G G / S G S V E S T E R F S
ORF RF [2]
<S N S I I M
<

210 220 230 240 250
GGCTGTTAACAGAAAGTTTACACAAGCTTGGAAAATAAGATAGGCTA
CCGACAAATTGTCTTCAAATGTGTTGAAACCCCTTTTATTCTATCCGAT
R Y E F M R T S F T O M A V L G R K S
ORF RF [2]
<

260 270 280 290 300
ACATTGGAATAAAACAAGGCTGGATTTATTATTCAGCCTTTAAATGT
TGTAACCTTATTTGTTCCGACCTAAATAAGGTGGAAAAATTACA

310 320 330 340 350
AAAATAAAAATACAGGGTTAAATAATCAAGTGTGCTGCGGATGAGAA
TTTTATTTATGTCCCATTATTAGTTACACAGCACCTACTCTT

360 370 380 390 400
GATAAAAACTATCTCTTAGAGAATAGGCCTCCTATTTATTATTAGGAG
CTATTTTGATAGAGAATCTCTTATCCGGAGGAGATAAAATAATCCTC
ORF RF [
<
410 420 430 440 450

1007410-27032003

Express Mail No. EL818623952US
Inventor(s): Cvitkovitch et al
Atty Dkt No.: 1889-00401
Contact: C.M. Mintz (713)238-8000
Sheet 10 of 19

AACGAACCTTATTACTACTAACGAAACAAACATTTGACCAAAACCTA
ORF RF [4] C

460 470 480 490 500
AATGTTCAAGAATATGATTCACCTTGTAAACCAATACCGCGATTGGAG
TTACAAGTTCTTATACTAAGTGGAAACGATTGGTTATGGCGCTAACCTC
< H E L I H M V K A L G I G R N S
ORF RF [4] C

510 520 530 540 550
CCTTTAGTGGAAATAGTTCTTTAAAAAATTTACTCACATCTATTTGTTT
GGAAATCACCTTATCAAACAAATTTAAAATGAGTGTAGATAAACAAA
< G W R T I S Y N E K P E T I C S V D O N S O N K
ORF RF [4] C

560 570 580 590 600
TTCTTTGGGAATTCTGAATGATAAACGACTATACTGCCATTTTCTTA
AAGAAACCACCTTAAGACTACTATTTCTGATATGACGGTAAAAGAATT
< E M R T S N Q I I F V U I S G N E K K F
ORF RF [4] C

610 620 630 640 650
AAAAGCTAACTGAATTTCAGGATTTAATGATTGAAAGCAGCCTCAATG
TTTCCGATTGACTTAAAGCTCTAAATTACTAAGCTTCGTCGGAGTTAC
M>

< F P A M L Q T N E R P N L S E F P A M E V E T
ORF RF [4] C

660 670 680 690 700
GCATTATCACACAAGATAGAAAGTATGGAATGAAATCAAGCAACTCCAT
CGTAATAGTGTGTTCTATCTTCTACACCATTACTTTAGTCGTTGAGGTA
A L S H K I E S M V M K S S N S I>
ORF RF [3]>
< A N D D C E L I S H T T I F G D E E E M
ORF RF [4] C

710 720 730 740 750
CTCAGGCAGTTGATTACTTGAGACTCTACATTGACAGCAATCTTT
GAGTCCGTCAACATAAAATGAACCTCTGAAGATGTAACGTGCGTAGAAAA
S G S C I L L E T S T L T A I F>
ORF RF [3]>
< E P T L Q T K V S S V E V N V A P I K K K
ORF RF [4] C

760 770 780 790 800
TATTCTGAGCTCTAACGATTTGCTGACAAGATAACCCCTGACAGCATCG
ATAAGACTCGAACGATTCTAAACGACTGTTCTATGGGAACTGTCGTAGC
L F>

>
< N Q V A E Y L T K A S M I T G K V V A D
ORF RF [4] C

810 820 830 840 850
TTTGAAATATTAGCTAGATGGCCGATATTATAACGGGTATCCTGCAATTG
AAAACCTTATAATCGATCTACCGGCTATAATATTGCCCATAGGACGTTAAC

00623017 - 040101

Express Mail No. EL818623952US
Inventor(s): Cvitkovitch et al
Atty Dkt No.: 1889-00401
Contact: C.M. Mintz (713)238-8000
Sheet 11 of 19

< D R O P T I V E A V L E C G I N E V Y C R S T D / O Q C S >
ORF RF[4] C

860 870 880 890 900

- ATGTCCTGTTTTCTAAGATTTGATGGTAAATCTTTCAATACTAGCTA
TACAGGACAAAAAAGATTCTAAACTACCAATTAGAAAAGTTATGATCGAT
< H G E T K E L I Q R V P E K V T S A L >
ORF RF[4] C

910 920 930 940 950

- AATCTTTATTTCAATGCCAACTCTGAGGCTAGTTAAAATATTCAAGATAA
TTAGAAATAAAAGTTACGGATTAGACTCCGATCAATTATAAGTCTATT
< D V K N E B D G E R H V S T M T E N T N Y >
ORF RF[4] C

960 970 980 990 1000

- TCATGGCGAAACTCGAATATCCTGTAAAGAGATTCTATTGCTGACT
AGTACCGCCTTGAAGCTTATAGAACATTCTCAAGATAAACGACTGA
< M A E T S N I L V K R F Y L L T >

< D H R E S R I D E K Y L S E P D Q O S >
ORF RF[4] C

1010 1020 1030 1040 1050

- ATACTGGGTGATATTCGAATCTGAGCTTCTTTGTGCCATTATCTCAT
TATGACCCACTATAAGCTTAGACTCGAAGGAAACACGGTAATAGAGTA
< I L G D I S N L S F L L C H Y L I >

< X O T I N R I Q A V E K Q R M T E N >
ORF RF[4] C

1060 1070 1080 1090 1100

- TTTGAAACCTTTGGTATATTGGCTAAAGAGATCAGAAATCAA
AAACTTGGAAAACAACCATATAACCGAATTCTACTCTAGTCTAGTTT
< L N L L F G I L A >

< Q V X Q K T Y O S L E F M S L E C I L >
ORF RF[4] C

1110 1120 1130 1140 1150

- AATAAAATAAGATAGACAATAACGACAAATTACGAAATTAAAGTCGG
TTATTCTATCTGTATTGCTGTTAAATGCTTTAAATTCAGCC
< F U L L Y V P V V B K R F K L T P >
ORF RF[4] C

1160 1170 1180 1190 1200

- TATCACATTATAACTCTCTATAACATACAATAACCTGTATTAAAAGGTAGT
ATAGTGTAAATTGAGAGATATTGTATGTTATGGACATAATTCCATCA
< T V M Y S E I V Y L V O N I L L Y Y >
ORF RF[4] C

1210 1220 1230 1240 1250

- ATAGAACCATAGTAATATTGATGGAAATCAAGCGTTTTGACCTTCATC
TATCTTCCGTATCATTATAAGTAACCTTAGTTCGAAAAAAACTGGAAGTAG
< S E L M T I N M R E Y D R M K K V Y K M >
ORF RF[4] C

Express Mail No. EL818623952US
Inventor(s): Cvitkovitch et al
Atty Dkt No.: 1889-00401
Contact: C.M. Mintz (713)238-8000
Sheet 12 of 19

1260 1270 1280 1290 1300
TTGGTCAAACATCTTTAAGTCGACCAATATCAACATTGAACACACTGAG
AACCGAGTTGATAGAAATTCAAGCTGGTTATAGTTGTAACCTGTGTGACTC
< K T L S D R E R G I D D V N E F V S L
ORF RF[4] C

1310 1320 1330 1340 1350
AAAGAGGTAACCTAGCGCTATTCCCTGCAAACTCGATCATATAGGTGGTTA
TTTCTCCATTGAATCGCGATAAGGACGTTGAGCTAGTATATCCACCAAT
< P E Y S I D A T G A P E P M M V E T T A Y
ORF RF[4] C

1360 1370 1380 1390 1400
TAATGCTACTGCCATTACAATTCCCTTGAGFTCCATCCAAGATAAAGAAT
ATTACGATGACGGGTAATGTTAAGGAACCTAAGGTAGGTCTATTTCTTA
< S S G M V I G O T G D L T P F
ORF RF[4] C

1410 1420 1430 1440 1450
ATGATTGCCGCTAAACAAGTCAGAACTGGCAACAGGCAGCAGACCATA
TACTAACGGCGGGATTGTTCAAGTCTTGACCGTTGTCGTCGTCTGGTAT
< T A A P R R E L V D D S H S A V M P T T E G M Y
ORF RF[4] C

1460 1470 1480 1490 1500
AAATATATTAGAGAAAGACTATTCTGCTATTAAGATAAATTGATAAAG
TTTATATAAATCTCTTCTGATAAGACAGATAATTCTATTTAACTATTTC
< P I N L S V L S N Q R N E L V M I S L A
ORF RF[4] C

1510 1520 1530 1540 1550
CTATAAAATAAGAGGCTCTGCAGGATAAAACAGGTTACGTTCAACCATC
GATATTCTTCTCCGAGACGTCTTATTTGTCAAATGCAAGTGGTAG
< P E Y L E B A P E V P E L T N V N V M
ORF RF[4] C

1560 1570 1580 1590 1600
GTAACAGCAATCATTATCAGAAAATTGCTTATCGAAAAAAGAGTTAATT
CATTGTCGTTAGTAATAGCTTTAAGAATAGCTTTCTCAATTAAAG
< T V A I M I L E N S I S F M L T Y D E
ORF RF[4] C

1610 1620 1630 1640 1650
CTTTTCGATAAAAGTGCACATTACTTACCTTAGAAAATAGAAACAGAGAA
GAAAAAGCTATTCACTGTAATGAATGGAATCTTTATCTTGTCTCTT
< K K S L T V W S V K S F L F L L F
ORF RF[4] C

1660 1670 1680 1690 1700
ATAGAACGGTTAGATAAGTAAATAACCATTTGAAAGTATCATTAAGGCT
TATCTTGCCAATCTATTCAATTATTTGCTAAACTTCAAGTAAATTCCGA
< I V T L Y T L E G M S L I X M S L A
ORF RF[4] C

Express Mail No. EL818623952US
Inventor(s): Cvitkovitch et al
Atty Dkt No.: 1889-00401
Contact: C.M. Mintz (713)238-8000
Sheet 13 of 19

1710 1720 1730 1740 1750
TCATTCACTTGTCTCCTTGATCAGCAATCACAGCTCTCAGTTTGTT
AGTAAGTAAACGAGAGGAAACTAGTCGTTAGTGTGAGAGTCAAAACAA
~~E E N M~~

<K A R R Q D A T V A P N G K T
ORF RF[5] C

1760 1770 1780 1790 1800
AACTTAGTCGTGAAATAAGACAAGACTTATTATTTCGAAAATAGGCCAA
TTGAATCAAGCRCTTATTCTGTTCTGAATAATAAGCTTTATCCGGTT
~~E L K T R S E L C S O N N R F Y A T E~~
<ORF RF[5] C

1810 1820 1830 1840 1850
GGCTTTTCCGATCAATACGGTAATATTGGCAGGATTGACAATAAAG
CGCAAAAAAGGCTAGTTATGCCATTATAACCGTCTTAACGTATTTC
~~E R A K R D Y T E N A V P N D V E P S~~
<ORF RF[5] C

1860 1870 1880 1890 1900
AGCGATGGCACTGAAAAGTCTCTTATCCATTAAACAATGTCAGTCATA
TCGCTACCGTGACTTTTCAGAGAATAGGTAAAATTGTTACAGTCAGTAT
M A L K K S L I H F N N V S H>
ORF RF[1]
~~E R H C Q F L R K D M X V I D S T M~~
<ORF RF[5] C

1910 1920 1930 1940 1950
CTGCCGTAGAATTCAATCCGTTCATCATAAGTATAAAGGCAGAGCTTATG
GACGGCATCTTAAGTTAGGCAAGTAGTATTCACTATTCCGTCTCGAACATAC
T A V E F N P F I I S I K A E L M>
ORF RF[1]
~~S E G A Y F E E I R E D Y X T A Y I C C L K H~~
<ORF RF[5] C
<A T S N L G N M M L I F A S S I
<ORF RF[6] C

1960 1970 1980 1990 2000
GGCTGTTGAACATGTTCAAATAACAGAAATCTCAGCAAAAGGGACCTGAA
CCGACAACCTCTACAAAGTTTATGTCTTAGAGTCGTTCCCTGGACTT
G C>
<A T S S T E P Y L D E A F P P V O P E
<ORF RF[5] C
<P Q Q L H K L I C F R L L L S R F
<ORF RF[6] C

2010 2020 2030 2040 2050
ACTGAGTTCAAGATGAATGGAAAATAAAAAGTCCCTGATTGACCATTG
TGACTCAAAGTCTACTTACCTTTATTTCAGGAGACTAACCTGGTAAG
~~S O T E S S H F I P L B E N S O G N~~
<ORF RF[5] C
<S L K L H I S F L F N R Q N V M R

< _____ ORF RF[6] C _____ >

2060 2070 2080 2090 2100

TTCTGGCTGTTCCATAGCATAATACAGCGCTGATTCAATGCCGTGGGA
AAGACCGACAAAAGGTATCGTATTATGTCGGACTAACGTTACGCCACCC
K S S E M A Y Y L B S G E R P H S

< _____ ORF RF[5] C _____ >

< R A T K W L M >

< _____ ORF RF[6] C _____ >

2110 2120 2130 2140 2150

GAACTCCTCAGGATTCAAAGATTTATCAATAAAATCCAAAAGCAGATACCT
CTTGAGGAGTCCTAACGTTCTAAATAGTTAGGTTCTATGGGA
C T T C T T C A C T C C C G T A T T G A G T C T T A C A G C A C T G T T C T G T

< _____ ORF RF[5] C _____ >

2160 2170 2180 2190 2200

GATACTGAAAAGTGAGGGGCATAAACTCAGAACATGTCGTGACAAGACA
CTATGACTTTCACTCCCCGTATTTGAGTCTTACACAGCACTGTTCTGT

M C R D K D >

< _____ ORF RF[5] C _____ >

2210 2220 2230 2240 2250

ATAACTGCACTAGGATTATGCTGCTTAATCTGATTGGCTACTTCCAGTCC
TATTGACGTGATCCTAACAGACAGATTAGACTAACCGATGAAGGTCAAGG
N N C T R I M L S N L I G Y F Q S >

< _____ ORF RF[5] C _____ >

2260 2270 2280 2290 2300

TTTCTTTCTCTTTGATTCAATATCCAAAAGAAAAATCTGGTGAT
AAAGAAAAGGAGAAAAAACTAAAGTATAGGTTTTCTTTAGACCACTA
F L F L P F D F N I Q K E N L V I >

< _____ ORF RF[5] C _____ >

2310 2320 2330 2340 2350

TGCCCTTTCAAGGATAGCTCAATAAGTTGTTGGTTCCAAAATA
ACGGAAAAGTCCCTATCGAGTTATCAACACACCAAAAGGTTTTAT
A L F R D S V N K L L W F S K N >

< _____ ORF RF[5] C _____ >

2360 2370 2380 2390 2400

GTCAATTCTTATAAGACCAATTCTTCTCATGATAGCTGCAATGGT
CAGTTAAGAAATATTCTGGTAAAAAAAGAAAAGTACTATCGACGTTACCA
S Q F F I R P I F F F H D S C N G >

M I A A M V >

< _____ ORF RF[5] C _____ >

TOP SECRET//NOFORN//COMINT

Express Mail No. EL818623952US
Inventor(s): Cvitkovitch et al
Atty Dkt No.: 1889-00401
Contact: C.M. Mintz (713)238-8000
Sheet 15 of 19

2410 2420 2430 2440 2450
GGTTCAAGACGTCCTTGTGTAaaaaATCATCTCCAATACAAAAATAG
CCAAAGTTCTGCAGGAACAACATTAGTAGAAGGTTATGTTTTATC
G F K T S L >

> V S R R P C C K K S S S N T K I >
< T D E G R G Q Q L P A D D E D A E E I N S
ORF RF [5] C >

2460 2470 2480 2490 2500
AAATCATTATTCTCCTTAAATCTTCTATTTAGGTTAGCTGATTAACACT
TTTACTAATAAGAGGAAATTAGAAGATAAAATCCAATCGACTAATTGTGA
E I I S P L I F Y L G >

< >
2510 2520 2530 2540 2550
ATACACAGAAAAGGTATAAACGATATCACTCAATAAAATCTACTAACTT
TATGTGTCTTTCCATATTGCTATAGTGAGTTATTTAGATGATTGAA
AATAACC
TTATTGG

Figure 10

The comX nucleotide sequence, amino acid sequence, and its local region with 100bp included both upstream and downstream (promoter is upstream).

➤ S. mutans comX gene

```
ATGGAAGAAGATTTGAAATTGTTTAATAAGGTTAACGCCATTGTATGAAATTAG  
CCGTTATTACTTATTAAAATGTGGACTCGTAAAGATTGGCAACAAGAGGGAAATGTTGA  
TTTGCACCAATTATTAAGGAACATCCAGAATTAGAAGAGGATGATAACAAATTGTAT  
ATCTATTAAAGACACGTTTCTAATTACATTAAAGATGTTTGCCTCAGCAAGAAAG  
TCAGAAACGTCGTTTAATAGAATGTCTATGAAGAAGTCGGTAGAGATTGAACACTGTT  
TGTCAAGTGGCGGTATGCAATTGGATGAATATATTATTTCGTGATAGTTGCTTGCA  
TATAAACAAAGGTCTGAGTACTGAAAAGCAAGAGCTGTTGAGCGCTGGTAGCAGGAGA  
GCACTTTTGGAAAGCAAAGTATGCTGAAAGATTACGTAAAAAATTAAAGTGAATTAA  
AGGAAAAA
```

➤ S. mutans ComX protein

```
MEEDFEIVFNKVVKPIVWKLSRYYFIKMWTREDWQQEGMLILHQLLREHPELEEDTKLY  
IYFKTRFSNYIKDVLRQQESQKRRFNRMSYEVEGEIEHCLSSGGMQLDEYILFRDSLAA  
YKQGLSTEKQELFERLVAGEHFLGRQSMLKDLRKKLSDFKEK
```

➤ S. mutans comX gene local region

```
GTAAATAAAACAGCCAGTTAACGATGGGACATTATGCTCTTAAAGTCTTTTCG  
TTTATAATAATTATTATTAAAGGAGGTATCGTAATAGATGAAAGAAGATTGAA  
ATTGTTTTAATAAGGTTAACCCAATTGTATGAAATTAGCCGTATTACTTTATTAA  
AATGTGGACTCGTAAGATGGCAACAAGAGGGAAATGTTGATTGACCAATTATTAA  
GGGAACATCCAGAATTAGAAGAGGATGATACAAATTGTATATCTATTAAAGACACGT  
TTTCTAATTACATTAAAGATGTTTGCCTCAGCAAGAAAGTCAGAAACGTCGTTTAA  
TAGAATGTTTATGAAAGAAGTCGGTGAGATTGAAACACTGTTGTCAGTGGCGGTATGC  
AATTGGATGAATATATTATTTCGTGATAGTTGCTTGCATATAAACAAAGGTCTGAGT  
ACTGAAAAGCAAGAGCTGTTGAGCGCTTGGTAGCAGGAGAGCACTTTTGGGAAGGCA  
AAAGTATGCTGAAAGATTACGTAAAAAATTAAAGTGAATTAAAGGAAAATAGTTAAAAA  
GGGAAAGAATGGAACATGTGATTGTACCATTCTTTGGTTGAAAATTAAAGAAAAGTTA  
TTATAAATTATTGGTTAACATGCCATATTA
```

09833012 " 0431004

Figure 11.

The comA and comB nucleotide and amino acid sequences. ComA and ComB are the components of the CSP exporter.

> S. mutans comA gene

```
ATGAAACAAGTTATTATGTGTTTAATCGTCATAGCCGTTAACATTCTCTTAGAGAT
TATCAAAAGAGTAACAAAAAGGGGAGGGACAGTTTCGTCATCTAACCTTACAGATG
GGCAGTCTAAGTTGGCGCAGACATTATAAGCTAGTACCTCAGATTGATACCAGA
GAATGTGGGCCGGCAGTGTGGCATCTGTTGCAAAGCATTACGGATCTAACATTACTCTAT
CGCTTATCTCGGGAACTCTCAAAGACTAACAGCAGGGAACACAGCTCTGGCATTG
TTGAAGCTGCTAAAAAGTTAGGCTTGAACACGGCTCATCAAGGGGATATGACGCTT
TTTGATTATAATGATTGACCTATCCTTTATCGTCATGTGATTAAAGGAAAACGTCT
GCAGCATTATTATGTCGTCTATGGCAGCCAGAATAATCAGCTGATTATGGAGATCCTG
ATCCTTCAGTTAAGGTGACTAGGATGAGTAAGGAACGGCTTCAATCAGAGTGGACAGGC
CTTGCATTTCTAGCTCTCAGCTTAACATATAAGCCTCATAAAAGGTGAAAAAAATGG
TTTGTCTAATTCTCCCGTGATCTTAAGCAGAAAGCTTGTGACTTATATTATCA
TAGCTAGCTTGATTGTGACGCTCATTGATATTGTCGGATCATACTATCTCCAAGGAATA
TTGGACGAGTACATTCTGATCAGCTGATTCAACTTTAGGAATGATTACGATTGGCT
GATAATAACCTATATTATCCAGCAGGTCACTGGCTTTGCAAAGAATAACCTCTGGCCG
TACTCAGTTGCGTTAGTCATTGATGTTATCCTGTCTTATATCAAACATATTTCAG
CTTCCTATGCTTTGCGACAAGGCAACAGGAGAAATCACGTCTCGTTACAGA
TGCCAATCAGATTATTGATGCTGAGCGTCAACCCTTCAATCTTGTAGATATGA
CTATGGTAATTGGTGGTGGGTTTGTGGCGAAAACAATAACCTTCTTCTA
ACCTTGCTCTCCATTCCGATTTATGCCATCATTATTTGCTTCTGAAACCCCTTGA
GAAAATGAATCACGAAGTGTGGAAAGCAATGCTGTTAGTTCTCTATCATTGAAG
ATATCAATGGATGAAACCAATTAAATCACTCACAAGTGAGTCCGCTCGTTATCAAAC
ATTGATAGTGAATTGTTGATTATTGGAGAAAACCTTAAGCTACACAAGTATAGTC
CATTCAAACCGCATTAAAAGCGGTGTAAGCTTATCCTCAATGTTGTCATCTCTGGT
ATGGCTCTCGTCTAGTTATGGATAATAAAATCTCAGTTGGTCAGCTTATCACCTTAAT
GCTTGCTGCTTATTCTCAAATCAAATTGAAAATATTATCAATCTGCAATCAAAC
GCAGTCAGCTCGCGTTGCCAATACACGTCTTAATGAGGTCTATCTGTCGAATCTGAAT
TTGAAAAAGACGGCGATTTACGAAAATAGCTTTAGATGGTGTGATATTGCTTGA
AAATCTTCTTATAAATATGGATTGGCGAGATAACCTTATCAGATATTAAATTATCAAT
CAAAAAGGCTCCAAGGTCACTAGTTGGAGCCAGTGGTTCTGGTAAAACAACCTTG
CTAAACTGATTGTCATTCTACGAGCTAACAAAGGGATTGTTGCAATCAATGGCAAT
GATTAAAAGTTATTGATAAGACAGCTTGCAGGGCATATTAGCTATTGCCGCAACA
GGCCTATGTTTAGTGGCTCTATTATGGATAATCTGTTTAGGAGCTAAAGAAGGAA
CGAGTCAGGAAGACATTACGAGCTGCTGTGAAATTGCTGAAATCCGCTCGGACATTGAA
CAAATGCCCTAGGGCTATCAGACAGAGTTATCAGATGGTGCCTGTTCTGGCGGTCA
AAAACAGCGGATTGCTTCTAGCTAGGGCTTATTAAACACAGGCACGGTTTGATTCTGG
ATGAAGCCACCAGCAGTCTGATATTGACAGAAAAGAAAATTATCAGCAATCTCTTA
CAGATGACGGAGAAAACAATAATTGTTGCCCACCGCTTAAGCATTACAGCGTAC
TGACGAAGTCATTGTCATGGATCAGGGAAAATTGTTGAACAAGGCACTCATAAGGAAC
TTTAGCTAAGCAAGGTTCTATTATAACCTGTTAAT
```

> S. mutans ComA protein

008304-1001

MKQVIYVVLIVIAVNILLEIIKRVTKRGTVSSSNPLPDGQSKLFWRRHYKLVPQIDTR
DCGPAVLASVAKHYGSNYSIAYLRELSTKNQGTTALGIVEAAKKLGFETRSIKADMTL
FDYNDLTYPFIVHVIKGKRLQHYYVYGSQNNQLIIGDPDPSVKVRMSKERFQSEWTG
LAIFLAPQPNYKPHGEKNGLSNFFPLIFKQKALMTYIIIASLIVTLIDIVGSYYLQGI
LDEYIPDQOLISTLGMITIGLIITYIIQQVMAFAKEYLLAVLSRLVIDVILSYIKHIFT
LPMSFFATRRTGEITSRFTDANQIIDAVASTIFSIFLDMTMVLVGGVLLAQNNNLFFL
TLLSIPYAIIFIAPFLKPFEKMHEVMESNAVVSSSIIEDINGMETIKSLTSESARYQN
IDSEFVDYLEKNFKLHKYSAIQTALKSGAKLILNVVILWYGSRLVMDNKISVGQLITFN
ALLSYFSNPIENIINLQSKLQSARVANTRLNEVYLVESEFEKDGDLSSENSFLDGDISFE
NLSYKYGFGRDTLS DINLSIKKGSKVSLVGASGSGKTTIAKLIIVNFYEPNKGIVRINGN
DLKVIDKTALRRHISYLPQQAYVSGSIMDNLVLGAKEGTQSEDIIRACEIAEIRSDIE
QMPQGYQTELSDGAGISGGQKQRIALARALLTQAPVLILDEATSSLDILTEKKIISNLL
QMTEKTIIFVAHRLSISORTDEVIVMDQGKIVEQGTHKELLAKQGFVYNLFN

➤ *S. mutans* comb gene

ATGGATCTAAATTTCACAAAGTCAGAATTATAGGAGACGCTATCATAATTTGC
GACACTATTAAATTGTTCTTGGTCTGATTATCTTCTTGGTCATATTCTTTGTT
TTGCTAAAAAGAAATTACAGTGATTCTACTCGTGAAGTTGCACCAACAAAGTTGTA
GATGTTATCCAATCTTACAGTGACAGTTCAATCATTAAAATAATTAGATAATAATGC
AGCTGTTGAGAAGGGAGACGTTTAATTGAATATTCAAGAAAATGCCAGTCAAACCGTC
AGACTGAACAAAAGAAATTATAAAAGAAAGACAAAACGAGAAGAGAAGGAAAAGAAA
AAACACCAAAAGAGCAAGAAAAAGAAGAACAGCTAAGAGCAAGAAAGCTTCAAAGATAAA
GAAAAAGAAATCGAAAGACAAGGAAAGCAGCTCTGACGATGAAATGAGACAAAAAAGG
TTTCGATTGGCTTCAAGAGATGGTATTATTACATACCAATCCCAAATATGATGGTGC
AATATTATTCCGAAGCAAACCGAGATGCTCAAATCTATCCTGATATTCAAACAAAG
AAAAGTGTAAATCACCTATTATGCTTCTGATGATGTTGTTCTATGAAAAAGGGC
AAACCGCTCGTCTTCCTTGGAAAAAAAGGGAAATGACAAGGTTGTTATTGAAGGAAA
ATTAACAATGTCGCTTCATCAGCAACTACTAAAGGAAATCTCTTAAAGGTTAC
TGCCAAAGTAAAGGTTCTAAGAAAATAGCAAACTCATCAAGTATGGTATGACAGGCA
AGACAGTCACTGTCATTGATAAAAGACTTATTTGATTATTCAAAGATAAATTACTG
CATAAAATGGATAAT

➤ *S. mutans* ComB protein

MDPKFLQSAEFYRRRYHNFTLLIVPLVCLIIIFLVIFLCFAKKEITVISTGEVAPTKVV
DVIQSYSDDSIIKMNLDNNAAVEKGDLIEYESENASPNRQTEQKNIikerQKREEKEKK
KHQKSKKKKSKSKASKDKKKSKDKESSSDDENETKVSIFASEDGIiHTNPKYDGA
NIIPKQTEIAQIYDPDIQKTRKVLITYYASSDDVSMKGQTARLSLEKKGNDKVVIEGK
INNVASSATTTKKGNLFKVTAKVKVSKNSKLICKYGMTGKTVTVIDKTYFDYFKDKLL
HKMDN

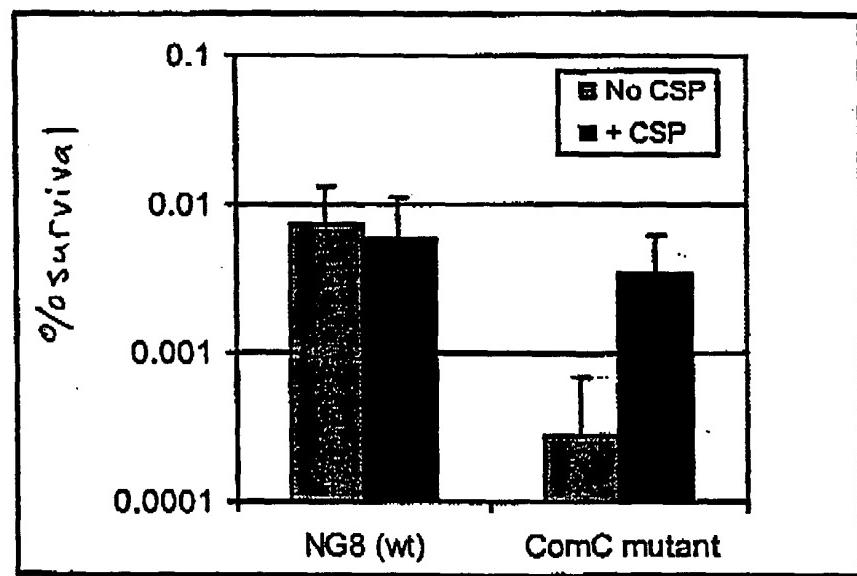


Figure 12

2025 RELEASE UNDER E.O. 14176